UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/420,509	10/18/1999	ALEXANDER FRANZ	80398.P282	2902

7590 01/31/2002

MARIA MCCORMACK SOBRINO
BLAKELY SOKOLOFF TAYLOR & ZAFMAN LLP
12400 WILSHIRE BOULEVARD
7TH FLOOR
LOS ANGELES, CA 90025

EXAMINER	
FLEURANTIN, JEAN B	
ART UNIT	PAPER NUMBER
2172	

DATE MAILED: 01/31/2002

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No. 09/420,509	Applicant(s) FRANZ ET AL.	
	Examiner Jean B. Fleurantin	Art Unit 2172	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____ .

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 1-88 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1-88 is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☐ The proposed drawing correction filed on _____ is: a)☐ approved b)☐ disapproved by the Examiner.

If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____ .

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

a) ☐ The translation of the foreign language provisional application has been received.

15)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☒ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449) Paper No(s) 2,4,6 .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

DETAILED ACTION

1. Claims 1-88 are presented for examination.

Information Disclosure Statement

2. The information Disclosure Statement, PTO-1449, have been fully considered.

Drawings

3. The Formal Drawings are required in response to this Office Action.

Claim Rejections-35 U.S.C. § 103 (a)

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1-88 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bahl et al. (U.S. Pat. No. 5,033,087) ("Bahl").

As per claims 1, 21, 41, and 61 Bahl substantially teaches a method for evaluating similarity among a plurality of data structures (thus, these tokens are analyzed to determine which word or words correspond to the sequence to the sequence of tokens, which is readable as evaluating similarity among a plurality of data structures) (see col. 1, lines 22-23), as claimed comprises analyzing each structure of said plurality of data structures to generate at least one substructure (thus, decision graph defines phonological rules which describe variations in the pronunciation of the various language components due to the context in which the component

occurs; which is readable as analyzing each structure of said plurality of data structures to generate at least one substructure) (see col. 2, lines 58-61); and

matching said at least one substructure to a database having a plurality of entries to obtain at least one matching entry (thus, after both the fast match operation and the detailed match operation the search processor 1020 invokes the language model 1010 to determine if the newly selected word fits in the context of the previously selected words, in addition to paring down the list of candidate words for application to the detailed match processor the language model 1010 distinguishes between the set of homophones provided as a result of the detailed match operation the language model used in the system shown in FIG. 8 is a three-gram language model, or stated otherwise a language model having statistics on the likelihood of occurrence of groups of three consecutive words; which is readable as matching said at least one substructure to a database having a plurality of entries to obtain at least one matching entry) (see col. 7, lines 56-64). But Bahl does not explicitly indicate the step of generating a match value using a relative entropy value corresponding to said at least one matching entry. However, Bahl implicitly indicates the step of the cluster number annotations of the set of sequences are used to determine the amount of disorder or entropy of the set and the context annotations are used to subdivide the set to produce subsets having less disorder than the parent set; which is readable as generating a match value using a relative entropy value corresponding to said at least one matching entry (see cols. 16-17, lines 66-3). Thus, it would have been obvious to a person of ordinary skill in the art at the time of the invention was made to modify the teachings of Bahl with the step of generating a

match value using a relative entropy value corresponding to said at least one matching entry. This modification would allow the teachings of Bahl to improve the accuracy and the reliability of the system and method to match linguistic structures using thesaurus information, and provide samples are then processed to develop a decision mechanism which relates the clusters to data indicating the contextual occurrence of the language components (see col. 2, lines 55-58).

As per claims 2, 22, 42, and 62 Bahl substantially teaches a method according as claimed, further comprises creating said plurality of entries in said database (thus, the acoustic vectors produced from a relatively large set of acoustic inputs are generated and stored in the cluster element 1110; which is readable as creating said plurality of entries in said database) (see col. 5, lines 39-41);

processing said plurality of entries in said database (thus, language model 1010 is used to determine which word is correct from a group of homophones the language model 1010 used in this embodiment of the invention determines which word of a group is the most likely based on the preceding two words derived by the speech recognition system the words determined by this language model analysis are the output of the speech recognition system; which is readable as processing said plurality of entries in said database) (see col. 4, lines 61-68).

As per claims 3, 23, 43, 63, and 82 Bahl substantially teaches a method according as claimed, wherein said creating further comprises creating said plurality of entries using a tool having a graphical user interface and exporting said plurality of entries to said database (thus, language model 1010 is used to determine which word is correct from a group of homophones

the language model 1010 used in this embodiment of the invention determines which word of a group is the most likely based on the preceding two words derived by the speech recognition system the words determined by this language model analysis are the output of the speech recognition system; which is readable as creating said plurality of entries using a tool having a graphical user interface) (see col. 4, lines 61-68).

As per claims 4, 24, 44, 64, and 83 in addition to the discussion in claim 1 above, Bahl further teaches the step of wherein said processing further comprises verifying said plurality of entries for validity (thus, if each of these stored vectors is considered to be a point in a state-space defined by a state vector of possible acoustic features, then the set of all points produced by the training data may be grouped into clusters of points in the state-space, each point in a given cluster represents a one centisecond sample of a vocal sounds which is statistically similar to the sounds represented by the other points in the cluster, each of the clusters in the state space may be thought of as a being representative samples of a probability distribution each of these probability distributions which may for example be Gaussian distributions defines a prototype for a label; which is readable as verifying said plurality of entries for validity) (see col. 5, lines 39-54).

As per claims 5, 25, 45, 65, and 84 Bahl substantially teaches a method according as claimed, wherein said processing further comprises storing said each entry of said plurality of entries together with said corresponding relative entropy value in a compressed format (thus, the step 1804 generates a question of the form x.sub.i .epsilon. S.sub.i which minimizes the

conditional entropy of the generating data that is marked "false" and which produces a net reduction in the entropy of the checking data, the algorithm used in the step 1804 is described below in reference to FIGS. 9A and 9B, if this question is "good" as determined, at step 1804, step 1806 causes it to be stored in the pylon at step 1810; which is readable as further comprises storing said each entry of said plurality of entries together with said corresponding relative entropy value in a compressed format) (see col. 15, lines 31-38).

As per claims 6, 26, 46, 66, and 85 Bahl substantially teaches a method according as claimed, further comprising extracting from a lexicon database having a plurality of elements each element associated to said each structure (thus, extracts all feneme sequences corresponding to individual phonemes in the training text these feneme sequences are grouped according to the phonemes they represent; which is readable as extracting from a lexicon database having a plurality of elements each element associated to said each structure) (see col. 9, lines 55-58), assigning at least one code of said each element to said each structure (thus, determines that the last leaf has been processed step 11016 is executed which stores all of the compound baseforms in a table indexed to the leaves of the decision tree; which is readable as assigning at least one code of said each element to said each structure) (see col. 19, lines 50-56), and retrieving said at least one code during matching to obtain said at least one matching entry (thus, invokes the subroutine NEXT LEAF to select the first leaf in the tree. Step 11006 then collects all feneme sequences that belong to the selected leaf, these feneme sequences are clustered at step 11008 using the same algorithm described above in reference to FIG. 7, assuming that the data used to

generate and check the decision tree includes approximately 3000 feneme sequences for each phoneme; which is readable as retrieving said at least one code during matching to obtain said at least one matching entry) (see col. 19, lines 24-32).

As per claims 7, 27, 47, and 67 Bahl substantially teaches a method according as claimed, further comprising reading lexical probability files and assigning a probability value to said each element of said plurality of elements in said lexicon database (thus, each of the transitions tr1 and tr2 has a transition probability and a vector of 200 probability values representing the probability that any of the 200 fenemes may be produced during the transition, the transition tr8 is a null transition; which is readable reading lexical probability files and assigning a probability value to said each element of said plurality of elements in said lexicon database) (see cols. 6-7, lines 65-3).

As per claims 8, 17, 28, 37, 48, 57, 68, 77, and 86 Bahl substantially teaches a method according as claimed, wherein each structure of said plurality of data structures is a representation of a linguistic expression (thus, by analyzing a training text and corresponding vocalizations, can generate a set of phonological rules, these rules are applied to a speech recognition system in the embodiment described below, they may also be applied to a speech synthesis system to change the pronunciation of a word depending on its context, or they may simply be analyzed by linguists to increase their knowledge of this arcane art; which is readable as wherein each structure of said plurality of data structures is a representation of a linguistic expression) (see cols. 3-4, lines 63-2).

As per claims 9, 18, 29, 38, 49, 58, 69, 78, and 87 Bahl substantially teaches a method according as claimed, wherein said database is a thesaurus hierarchy including a root entry, said plurality of entries depending from said root entry (thus, these words are arranged in a tree structure for use by the processor 1006, so that words having common initial phonemes have common paths through the tree until they are differentiated; which is readable as wherein said database is a thesaurus hierarchy including a root entry, said plurality of entries depending from said root entry) (see col. 6, lines 33-37).

As per claims 10, 19, 30, 39, 50, 59, 70, 79, and 88 the limitations of claims 10, 19, 30, 39, 50, 59, 70, 79, and 88 are rejected in the analysis of claim 1 above, and these claims are rejected on that basis.

As per claims 11, 20, 31, 40, 51, 60, 71, and 80 Bahl substantially teaches a method according as claimed, wherein said each element in said lexicon database is a word (thus, each word in a dictionary is represented as a sequence of phonemes, which is readable as wherein said each element in said lexicon database is a word) (see col. 6, lines 23-25).

As per claims 12, 32, 52, and 72, in addition to the discussion in claim 1 above, Bahl teaches the step of creating a plurality of entries in a database (thus, the acoustic vectors produced from a relatively large set of acoustic inputs are generated and stored in the cluster element 1110; which is readable as creating a plurality of entries in a database) (see col. 5, lines 39-41).

As per claims 13, 33, 53, and 73 Bahl substantially teaches a method according as claimed, further comprising storing said each entry of said plurality of entries together with said corresponding relative entropy value in a compressed format (thus, the step 1804 generates a question of the form x.sub.i .epsilon. S.sub.i which minimizes the conditional entropy of the generating data that is marked "false" and which produces a net reduction in the entropy of the checking data, the algorithm used in the step 1804 is described below in reference to FIGS. 9A and 9B, if this question is "good" as determined, at step 1804, step 1806 causes it to be stored in the pylon at step 1810; which is readable as further comprises storing said each entry of said plurality of entries together with said corresponding relative entropy value in a compressed format) (see col. 15, lines 31-38).

As per claims 14, 34, 54, and 74 Bahl substantially teaches a method according as claimed, further comprises creating said plurality of entries using a tool having a graphical user interface (thus, determines which word of a group is the most likely based on the preceding two words derived by the speech recognition system the words determined by this language model analysis are the output of the speech recognition system; which is readable as creating said plurality of entries using a tool having a graphical user interface) (see col. 4, lines 61-68);

exporting said plurality of entries to said database (thus, by replacing the each of the 200 value label probability vectors associated with the various transitions in the word model by a single 200 value probability vector, each element in this vector is the largest corresponding value

in all of the vectors used in the model; which is readable as exporting said plurality of entries to said database) (see col. 7, lines 48-53).

As per claims 15, 35, 55, and 75 in addition to the discussion in claims 1 and 12, Bahl further teaches analyzing each structure of said plurality of data structures to generate at least one substructure (thus, decision graph defines phonological rules which describe variations in the pronunciation of the various language components due to the context in which the component occurs; which is readable as analyzing each structure of said plurality of data structures to generate at least one substructure) (see col. 2, lines 58-61); and

matching said at least one substructure to a database having a plurality of entries to obtain at least one matching entry (thus, after both the fast match operation and the detailed match operation the search processor 1020 invokes the language model 1010 to determine if the newly selected word fits in the context of the previously selected words, in addition to paring down the list of candidate words for application to the detailed match processor the language model 1010 distinguishes between the set of homophones provided as a result of the detailed match operation the language model used in the system shown in FIG. 8 is a three-gram language model, or stated otherwise a language model having statistics on the likelihood of occurrence of groups of three consecutive words; which is readable as matching said at least one substructure to a database having a plurality of entries to obtain at least one matching entry) (see col. 7, lines 56-64).

As per claims 16, 36, 56, and 76, in addition to the discussion in claim 6 above, Bahl teaches verifying said plurality of entries for validity (thus, the feature selection element 1108

combines selected values of the vector signal SA to generate a vector, AF of acoustic feature signals; which is readable as verifying said plurality of entries for validity) (see col. 5, lines 30-32);

reading lexical probability files (thus, each transition has a probability associated with it and, in addition each of these transitions except the ones indicated by broken lines 'i.e. tr11, tr12, and tr13' has associated with it a vector of 200 probability values representing the probability that each of the respective 200 possible labels occurs at the transition, the broken-line transitions represent transitions from one state to another in which no label is produced; which is readable as reading lexical probability files) (see cols. 6-7, lines 63-3);

assigning a probability value to said each element of said plurality of elements in said lexicon database (thus, samples of a probability distribution each of these probability distributions which may, for example be assumed to be Gaussian distributions defines a prototype for a label or feneme, when the acoustic processor 1004 is in its training mode the cluster element provides the clusters to the prototype element which fits a Gaussian distribution to each cluster defining a prototype label which represents all points in the cluster, when the acoustic processor is in its labeling mode these prototypes are used by the labeller 1114 to assign labels to the feature vectors produced by the feature selection element 1108; which is readable as assigning a probability value to said each element of said plurality of elements in said lexicon database) (see col. 5, lines 50-62).

As per claim 81, in addition to the discussion in claim 1 above, Bahl teaches the step of a database having a plurality of entries (thus, the acoustic vectors produced from a relatively large set of acoustic inputs are generated and stored in the cluster element 1110; which is readable as a database having a plurality of entries) (see col. 5, lines 39-41).

5. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Kohler US Patent Number 6,212,500 relates to similarities between the different languages are exploited. Howard et al. US Patent Number 6,230,153 relates to log data by reducing large numbers of candidate rules to smaller rule sets; and also column 11, lines 20 through 32 Howard teaches relative entropy is a measure of expected information content for discriminating between two distribution. Horiguchi et al. US Patent Number 6,243,669 relates to speech or voice translation systems. Horiguchi et al. US Patent Number 6,330,530 relates to language translation systems.

Conclusion

6. Any inquiry concerning this communication from examiner should be directed to Jean Bolte Fleurantin at (703) 308-6718. The examiner can normally be reached on Monday through Friday from 7:30 A.M. to 6:00 P.M.

If any attempt to reach the examiner by telephone is unsuccessful, the examiner's supervisor, Mrs. KIM VU can be reached at **(703) 305-8449.** The FAX phone numbers for the Group 2100 Customer Service Center are: ***After Final*** **(703) 746-7238**, ***Official*** **(703) 746-7239**,

and ***Non-Official (703) 746-7240***. NOTE: Documents transmitted by facsimile will be entered as official documents on the file wrapper unless clearly marked "***DRAFT***".

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the Group 2100 Customer Service Center receptionist whose telephone numbers are **(703) 306-5631, (703) 306-5632, (703) 306-5633**.

Jean Bolte Fleurantin

January 24, 2002

JBF/

KIM VU
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100